

## CUBATURE ON WIENER SPACE IN INFINITE DIMENSION

CHRISTIAN BAYER AND JOSEF TEICHMANN

ABSTRACT. We prove a stochastic Taylor expansion for SPDEs and apply this result to obtain cubature methods, i.e. high order weak approximation schemes for SPDEs, in the spirit of T. Lyons and N. Victoir. We can prove a high-order weak convergence for well-defined classes of test functions if the process starts at sufficiently regular initial values. We can also derive analogous results in the presence of Lévy processes of finite type, here the results seem to be new even in finite dimension. Several numerical examples are added.

## 1. INTRODUCTION

Let  $H$  be a real separable Hilbert space. We consider the stochastic (partial) differential equation for a diffusion process  $X$  with values in  $H$

$$(1) \quad dX_t^x = (AX_t^x + \alpha(X_t^x))dt + \sum_{i=1}^d \beta_i(X_t^x)dB_t^i$$

or – in the presence of jumps – the stochastic differential equation for a jump-diffusion process  $X$  with values in  $H$

$$(2) \quad dX_t^x = (AX_{t-}^x + \alpha(X_{t-}^x))dt + \sum_{i=1}^d \beta_i(X_{t-}^x)dB_t^i + \sum_{j=1}^e \delta_j(X_{t-}^x)dL_t^j,$$

where, in general,  $A : \mathcal{D}(A) \subset H \rightarrow H$  denotes an unbounded linear operator,  $\alpha, \beta_1, \dots, \beta_d, \delta_1, \dots, \delta_e : H \rightarrow H$  denote  $C^\infty$ -bounded vector fields – i.e. the vector fields are smooth and all the derivatives (of degree  $\geq 1$ ) are bounded – and  $(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^d)_{t \geq 0}$  denotes a finite dimensional Brownian motion on the Wiener space  $(\Omega, \mathcal{F}, P)$ , and  $(L_t^j)_{t \geq 0}$  a compound Poisson process with jump-rate  $\mu_j$  for  $j = 1, \dots, e$ . The initial value  $x \in H$  appears as superscript in the notation  $X_t^x$  of the solution process, i.e.  $X_0^x = x$ . We assume that  $A$  is the generator of a  $C_0$ -semigroup denoted by  $(S_t)_{t \geq 0}$ . The main reference for equations of the type (1) is the monograph of G. da Prato and J. Zabczyk [5]. In the (general) Lévy case we refer to [6] and [19], even though we do not need results of their strength in our case since all Lévy processes under consideration are of finite type.

Due to the unboundedness of the operator  $A$  several concepts of solutions to (2) arise. We refer for the precise definitions to the excellent monograph [19]. The most direct analogue to the finite dimensional setting is the concept of a *strong solution*, which is defined by

$$(3) \quad X_t^x = x + \int_0^t (AX_{s-}^x + \alpha(X_{s-}^x))ds + \sum_{i=1}^d \int_0^t \beta_i(X_{s-}^x)dB_s^i + \sum_{j=1}^e \int_0^t \delta_j(X_{s-}^x)dL_s^j,$$

---

*Key words and phrases.* 60H10, 60H35.

The authors gratefully acknowledge the support from the FWF-grant Y328.

i.e. by the integrated version of equation (1). Besides the obvious integrability assumptions we need in particular that  $X_t^x \in \mathcal{D}(A)$  for all  $t \geq 0$  almost surely with respect to  $P$ . More relevant is the concept of a *mild solution*, which is related to (2) via variation of constants. Indeed, a mild solution is a process  $X_t^x$  satisfying

$$(4) \quad X_t^x = S_t x + \int_0^t S_{t-s} \alpha(X_{s-}^x) ds + \sum_{i=1}^d \int_0^t S_{t-s} \beta_i(X_{s-}^x) dB_s^i + \sum_{j=1}^e \int_0^t S_{t-s} \delta_j(X_{s-}^x) dL_s^j$$

given the obvious integrability assumptions. By Itô's formula, any strong solution is a mild solution, but not the other way round. Note that mild solutions need not be semi-martingales, because there is no semi-martingale decomposition if the process evolves outside of  $\mathcal{D}(A)$ . Consequently, a Stratonovich formulation of (2) is, in general, not possible for mild solutions.

Of course, neither strong nor mild solutions can usually be given explicitly, which makes numerical approximation necessary. We are interested in *weak approximation* of the solution in the sense that we want to approximate the value

$$P_t f(x) = E(f(X_t^x))$$

for a suitable class of test functions  $f : H \rightarrow \mathbb{R}$  at initial values  $x \in H$ . It is well-known that the function  $(t, x) \mapsto P_t f(x)$  solves the Kolmogorov equation in the weak sense, see for instance [5] in the diffusion case.

Let us assume for a moment that there are no jump-components: usually, infinite dimensional SDEs are numerically solved by finite element or finite difference schemes, see, for instance, [13], [23], [9] and [10]. For HJM models of financial mathematics a finite difference scheme and a finite element scheme have been implemented in [3]. This means that the original equation is projected onto some finite dimensional subspace  $H_h \subset H$  and  $A$  is approximated by some operator  $A_h$  defined thereon. This procedure, which corresponds to a space discretization of the stochastic PDE, is followed by a conversion of the stochastic differential equation on  $H_h$  to a stochastic difference equation by discretizing in time, using an Euler method or a related scheme. Finally, the stochastic difference equation is solved by Monte-Carlo simulation, which may be interpreted as a discretization on the Wiener space. For general information about approximation of finite dimensional SDEs see [14].

We want to tackle the problem in the reverse order: we want to do the discretization on the Wiener space  $\Omega$  first, reducing the problem to a deterministic problem, i.e., one replaces the  $d$ -dimensional Brownian motion with finitely many trajectories of bounded variation chosen with well-defined probabilities such that certain moments (of iterated Stratonovich integrals) match. The resulting deterministic problem can be solved by standard methods for the numerical treatment of deterministic PDEs, e.g. by standard finite element or finite difference methods. The benefit is that once the discretization on the Wiener space has been done, we can use the well-established theory of the corresponding deterministic PDE-problems, without any complications from stochasticity. Our method of choice for discretization on  $\Omega$  is “*cubature on Wiener space*”, developed by Terry Lyons and Nicolas Victoir in [18] and by Shigeo Kusuoka in [16] and [17], see also [22]. In the spirit of these methods we shall obtain weak approximation schemes of any prescribed order of convergence. Notice here that we discretize in the presence of the unbounded operator  $A$  in the drift vector field. Certainly our Assumptions 2.1 seem very restrictive, but these assumptions are the appropriate analogues of the assumptions

in finite dimension that the vector fields are bounded,  $C^\infty$ -bounded (see [17] and [18]).

Before going into details, let us motivate the use of cubature formulas in the present context. Let  $X_t^x(\omega)$  denote the solution of (1), formally rewritten in Stratonovich form, if each “ $dB_t^{i\prime\prime}$ ” is replaced by “ $d\omega^i(t)$ ”, i.e.

$$(5) \quad dX_t^x(\omega) = (AX_t^x(\omega) + \alpha(X_t^x(\omega)) - \frac{1}{2} \sum_{i=1}^d D\beta_i(X_t^x(\omega)) \cdot \beta_i(X_t^x(\omega)))dt + \\ + \sum_{i=1}^d \beta_i(X_t^x(\omega))d\omega^i(t)$$

for a curve function  $\omega = (\omega^1, \dots, \omega^d) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$  of bounded variation. Roughly speaking the idea of cubature on Wiener space is to construct short-time asymptotics (for some given degree of accuracy  $m \geq 2$ )

$$E(f(X_t^x)) = P_t f(x) = \sum_{l=1}^N \lambda_l f(X_t^x(\omega_l)) + \mathcal{O}(t^{\frac{m+1}{2}}),$$

with some positive, time-independent weights  $\lambda_1, \dots, \lambda_N$  satisfying  $\lambda_1 + \dots + \lambda_N = 1$  and some  $d$ -dimensional paths  $\omega_1, \dots, \omega_N$  of bounded variation. Of course, the weights and paths are chosen in a specific way, which will be described later in more detail, and the asymptotics will only hold for some class of test functions  $f$ . Notice in particular that the cubature paths  $\omega_1, \dots, \omega_N$  depend on the interval  $[0, t]$  – they become rougher as  $t$  approaches 0. The aforementioned procedure replaces the SDE by  $N$  deterministic, well-defined PDEs, which have unique mild solutions. The iteration of the short-time asymptotics due to the Markov property then yields a weak, high order approximation scheme.

*Here also the main advantage of cubature methods in contrast to Taylor methods gets visible.* The time-discretization in the realm of cubature methods always leads to reasonable expressions, namely to *reasonable partial differential equations of type* (5). If we wanted to apply the usual discretization methods in time like the Euler-Maruyama method, we might run into problems. Indeed, the naive Euler scheme is well-suited for the differential formulation (1) of the problem,

$$X_0 = x \text{ and } X_n = (AX_{n-1} + \alpha(X_{n-1}))\frac{t}{n} + \sum_{i=1}^d \beta_i(X_{n-1})\Delta_n B^i,$$

for  $n \geq 1$ , but it might immediately lead to some  $X_n \notin \mathcal{D}(A)$ . Even in the case of a well-defined strong solution, there is no reason why the discrete approximation should always stay in  $\mathcal{D}(A)$ . Hence the naive implementation of the Euler-Maruyama method does not work.

Only formulation (4) seems to be suitable for using an Euler-like method. If one understands the semigroup  $S$  well, one can approximate  $X_t^x$  by expressing the integrals in (4) as Riemannian sums, involving evaluations of  $S_{t-s}$ , which yields a sort of strong Euler method (see for instance the very interesting book [20] for strong convergence theorems in this direction).

We do not discretize the integral in formulation (4), but (weakly) approximate the Brownian motion's paths by paths of bounded variation such that we obtain a weak approximations of  $X_t^x$ .

The presence of jumps does not lead to more complicated expressions, since the short time asymptotics of a jump-diffusion can be easily derived from a diffusion's short-time asymptotics by conditioning on the jumps. The arising picture is the following. Discretizing the equation (2) means to allow a certain number of jumps

between to consecutive points in the time grid. Between two jumps we apply a diffusion cubature formula to express the short-time asymptotics. This yields as a corollary of the diffusion theory also the jump diffusion theory.

**Remark 1.1.** A direct application of the cubature on Wiener space technique to jump diffusions driven by Lévy processes of infinite activity is not possible. However, notice that any Lévy process can be approximated by processes with finite activity in a weak sense, see, for instance, [4] and [19], and therefore the solutions of the corresponding Hilbert-space valued SDEs converge weakly. In this sense, cubature methods can be use for jump diffusions driven by Lévy processes of infinite activity, too.

The article is organized as follows. In Section 2 we describe the analytic setting for a stochastic Taylor expansion to work. This is a delicate question since we deal with one unbounded vector field. In Section 3 and Section 4 we work out the cubature method from the scratch and prove the relevant convergence results in the diffusion case. In Section 5 we allow for jumps and prove the associated short-time asymptotics which is relevant to set up a weak approximation scheme. In Section 6 we apply our method to several examples to demonstrate the results.

## 2. SETTING AND ASSUMPTIONS

Let  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \geq 0})$  be a filtered probability space with a filtration  $(\mathcal{F}_t)_{t \geq 0}$  satisfying the usual conditions. Let  $(B_t)_{t \geq 0}$  be a  $d$ -dimensional Brownian motion and  $(L_t^j)_{t \geq 0}$ ,  $j = 1, \dots, e$  be  $e$  independent compound Poisson processes given by

$$L_t^j := \sum_{k=1}^{N_t^j} Z_k^j,$$

where  $N_t^j$  denotes a Poisson process with jump rate  $\mu_j > 0$  and  $Z^j = (Z_k^j)_{k \geq 1}$  is an i.i.d. sequence of random variables with distribution  $\nu_j$  for  $j = 1, \dots, e$ , such that each  $\nu_j$  admits all moments. All notions are with respect to the given filtration. We assume furthermore that all sources of randomness are mutually independent.

Let  $H$  be a separable Hilbert space. We furthermore fix a strongly continuous semi-group  $S$  on  $H$  with generator  $A$ . Let  $\alpha, \beta_1, \dots, \beta_d$ , the diffusion vector fields, and  $\delta_1, \dots, \delta_e$ , the jump vector fields, be  $C^\infty$ -bounded on  $H$ , that is, the vector fields are infinitely often differentiable with bounded partial derivatives of all proper orders  $n \geq 1$ . We consider the mild càdlàg solution  $(X_t^x)_{t \geq 0}$  of a stochastic differential equation

$$(6) \quad dX_t^x = (AX_t^x + \alpha(X_t^x))dt + \sum_{i=1}^d \beta_i(X_t^x)dB_t^i + \sum_{j=1}^e \delta_j(X_t^x)dL_t^j,$$

$$(7) \quad X_0^x = x \in H.$$

See [6] for all necessary details on existence and uniqueness for the previous equation. Notice furthermore the decomposition theorem in [8], which states that we do not need any further existence and uniqueness results in this case: in particular, we do not need to impose further (contractivity) conditions on  $A$  as in [6] in the finite activity case.

The previous conditions are slightly more than standard for existence and uniqueness of mild solutions, i.e. in [6] the authors need Lipschitz conditions on the vector fields, whereas we assume them to be  $C^\infty$ -bounded. In order to formulate a stochastic Taylor expansion we shall need one main assumption, which we formulate in the sequel. This assumption has already been successfully applied in several circumstances, e.g. [2], [7] or the recent [8].

We apply the following notations for Hilbert spaces  $\mathcal{D}(A^k)$ ,

$$\begin{aligned} \mathcal{D}(A^0) &:= H, \\ \mathcal{D}(A^k) &:= \{h \in H \mid h \in \mathcal{D}(A^{k-1}) \text{ and } A^{k-1}h \in \mathcal{D}(A)\}, \\ \|h\|_{\mathcal{D}(A^k)}^2 &:= \sum_{i=0}^k \|A^i h\|^2, \\ \mathcal{D}(A^\infty) &= \bigcap_{k \geq 0} \mathcal{D}(A^k), \end{aligned}$$

which we need in order to specify the main analytic condition for our considerations:

**Assumption 2.1.** We assume that  $\alpha, \beta_1, \dots, \beta_d$ , the diffusion vector fields, and  $\delta_1, \dots, \delta_e$ , the jump vector fields, map  $\mathcal{D}(A^k) \rightarrow \mathcal{D}(A^k)$  and are  $C^\infty$ -bounded thereon for each  $k \geq 0$ , that is, the vector fields are infinitely often differentiable with bounded partial derivatives of all orders  $n \geq 1$  on the Hilbert space  $\mathcal{D}(A^k)$  for each  $k \geq 0$ .

**Remark 2.2.** These assumptions are the appropriate analogues of the assumptions in finite dimension that the vector fields are bounded,  $C^\infty$ -bounded (see [17] and [18]). In order to establish a true convergence rate one needs an additional cut-off argument, which is outlined in Remark 4.9. This can – like in the finite dimensional setting – certainly be improved.

**Example 2.3.** In order to show examples of vector fields which are  $C^\infty$ -bounded on  $\mathcal{D}(A^k)$  consider the following structure. Let  $H$  be a separable Hilbert space and  $A$  the generator of a strongly continuous semigroup. We know that  $\mathcal{D}(A^\infty)$  is a Fréchet space and an injective limit of the Hilbert spaces  $\mathcal{D}(A^k)$  for  $k \geq 0$ . Following the analysis as developed in [7] (see also [15] and [12] where the analytic concepts have been originally developed), we can consider the vector field  $V : U \subset H \rightarrow \mathcal{D}(A^\infty)$ . If  $V$  is smooth in the sense explained in [7] and has the property that its derivatives of order  $n \geq 1$  are bounded on  $U \subset H$ , then  $V$  is obviously a  $C^\infty$ -bounded vector field and additionally  $V|_{\mathcal{D}(A^\infty)}$  is a Banach-map-vector field in the sense of [7]. Such vector fields constitute a class, where the above assumptions can be readily checked.

### 3. THE CASE WHEN $A$ IS BOUNDED LINEAR

We shall assume in this section that there are no jumps, i.e. we consider

$$(8) \quad dX_t^x = (AX_t^x + \alpha(X_t^x))dt + \sum_{i=1}^d \beta_i(X_t^x)dB_t^i.$$

In order to give an introduction to cubature on Wiener space, we consider the problem for a bounded operator  $A$ . In this case, there are virtually no differences to the finite dimensional setting, except the fact that the drift vector field does not need to be bounded by some constant on the whole Hilbert space (due to the presence of one linear operator in it). Remember that in [18] and [16] and [17] one deals with globally bounded vector fields. We shall circumvent this problem by a small refinement of the arguments.

Since mild and strong solutions coincide, we can always work with strong solutions, which are semi-martingales. Consequently, we can rewrite (8) into its Stratonovich form

$$(9) \quad dX_t^x = \beta_0(X_t^x)dt + \sum_{i=1}^d \beta_i(X_t^x) \circ dB_t^i,$$

where  $\beta_0 : H \rightarrow H$  denotes the Stratonovich-corrected drift, i.e.

$$(10) \quad \beta_0(x) = Ax + \alpha(x) - \frac{1}{2} \sum_{i=1}^d D\beta_i(x) \cdot \beta_i(x),$$

where

$$DF(x) \cdot v = \left. \frac{\partial}{\partial \epsilon} \right|_{\epsilon=0} F(x + \epsilon v)$$

denotes the Fréchet derivative of a function or vector field  $F$ . This notation enables us to write

$$dX_t^x = \sum_{i=0}^d \beta_i(X_t^x) \circ dB_t^i,$$

where we use the convention that “ $\circ dB_t^0 = dt$ ”.

The following notions form the core of cubature on Wiener space. We only give a short description and refer the reader to [18] and [22] for more details. Let  $\mathcal{A}$  denote the set of all multi-indices in  $\{0, \dots, d\}$ . We define a degree on  $\mathcal{A}$  by setting

$$\deg(i_1, \dots, i_k) = k + \#\{1 \leq j \leq k \mid i_j = 0\},$$

$k \in \mathbb{N}$ ,  $(i_1, \dots, i_k) \in \mathcal{A}$ . We have to count all the zeros twice because of the different scalings for  $t = B_t^0$  and the Brownian motion.

Recall that any vector field  $\beta$  can be interpreted as a first order differential operator on test functions  $f$  by

$$(\beta f)(x) = Df(x) \cdot \beta(x), \quad x \in H.$$

For a multi-index  $(i_1, \dots, i_k) \in \mathcal{A}$ ,  $k \in \mathbb{N}$ , let

$$B_t^{(i_1, \dots, i_k)} = \int_{0 \leq t_1 \leq \dots \leq t_k \leq t} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k}$$

denote the corresponding iterated Stratonovich integral. The iterated integrals form the building blocks of the stochastic Taylor formula, see [1].

**Proposition 3.1** (Stochastic Taylor expansion). *Let  $f \in C^\infty(H; \mathbb{R})$  and fix  $0 < t < 1$ ,  $m \in \mathbb{N}$ ,  $m \geq 1$ ,  $x \in H$ . Then we have*

$$f(X_t^x) = \sum_{\substack{(i_1, \dots, i_k) \in \mathcal{A} \\ \deg(i_1, \dots, i_k) \leq m}} (\beta_{i_1} \cdots \beta_{i_k} f)(x) B_t^{(i_1, \dots, i_k)} + R_m(t, f, x),$$

with

$$\sup_{x \in H} \sqrt{E(R_m(t, f, x)^2)} \leq Ct^{\frac{m+1}{2}} \max_{m < \deg(i_1, \dots, i_k) \leq m+2} \sup_{y \in H} |\beta_{i_1} \cdots \beta_{i_k} f(y)|.$$

**Remark 3.2.** Notice that under the assumptions of Proposition 3.1 the bound for the remainder term can be infinite. However, if  $f$  is bounded,  $C^\infty$ -bounded and has bounded support, then we can guarantee that

$$\sup_{y \in H} |\beta_{i_1} \cdots \beta_{i_k} f(y)| < \infty.$$

In the unbounded case this question is more subtle, see Example 4.8 and Remark 4.9.

Proposition 3.1 shows that iterated Stratonovich integrals play the rôle of polynomials in the stochastic Taylor expansion. Consequently, it is natural to use them in order to define cubature formulas. Let  $C_{bv}([0, t]; \mathbb{R}^d)$  denote the space of paths of bounded variation taking values in  $\mathbb{R}^d$ . As for the Brownian motion, we append a component  $\omega^0(t) = t$  for any  $\omega \in C_{bv}([0, t]; \mathbb{R}^d)$ . Furthermore, we establish the following convention: whenever  $X_t^x$  is the solution to some stochastic differential equation of type (1) driven by  $B$ , whether on a finite or infinite dimensional space,

and  $\omega \in C_{bw}([0, t]; \mathbb{R}^d)$ , we denote by  $X_t^x(\omega)$  the solution of the deterministic differential equation given by formally replacing all occurrences of “ $\circ dB_s^i$ ” with “ $d\omega^i(s)$ ” (with the same initial values), see equation (5) as compared to (1). Note that it is necessary that the SDE for  $X$  is formally formulated in the Stratonovich sense (recall that the Stratonovich formulation does not necessarily make sense for (1)).

**Definition 3.3.** Fix  $t > 0$  and  $m \geq 1$ . Positive weights  $\lambda_1, \dots, \lambda_N$  summing up to 1 and paths  $\omega_1, \dots, \omega_N \in C_{bw}([0, t]; \mathbb{R}^d)$  form a *cubature formula on Wiener space* of degree  $m$  if for all multi-indices  $(i_1, \dots, i_k) \in \mathcal{A}$  with  $\deg(i_1, \dots, i_k) \leq m$ ,  $k \in \mathbb{N}$ , we have that

$$E(B_t^{(i_1, \dots, i_k)}) = \sum_{l=1}^N \lambda_l B_t^{(i_1, \dots, i_k)}(\omega_l),$$

where we used the convention in line with the previous one, namely

$$B_t^{(i_1, \dots, i_k)}(\omega) = \int_{0 \leq t_1 \leq \dots \leq t_k \leq t} d\omega^{i_1}(t_1) \cdots d\omega^{i_k}(t_k).$$

Lyons and Victoir [18] show the existence of cubature formulas on Wiener space for any  $d$  and size  $N \leq \#\{I \in \mathcal{A} \mid \deg(I) \leq m\}$  by applying Chakalov’s theorem on cubature formulas and Chow’s theorem for nilpotent Lie groups. Moreover, due to the scaling properties of Brownian motion (and its iterated Stratonovich integrals), i.e.

$$B_t^{(i_1, \dots, i_k)} =_{\text{law}} \sqrt{t}^{\deg(i_1, \dots, i_k)} B_1^{(i_1, \dots, i_k)},$$

it is sufficient to construct cubature paths for  $t = 1$ .

**Assumption 3.4.** Once and for all, we fix one cubature formula  $\tilde{\omega}_1, \dots, \tilde{\omega}_N$  with weights  $\lambda_1, \dots, \lambda_N$  of degree  $m \geq 1$  on the interval  $[0, 1]$ . By abuse of notation, for any  $t > 0$ , we will denote  $\omega_l(s) = \tilde{\omega}_l(s/\sqrt{t})$ ,  $s \in [0, t]$ ,  $l = 1, \dots, N$ , the corresponding cubature formula for  $[0, t]$ .

**Example 3.5.** For  $d = 1$  Brownian motions, a cubature formula on Wiener space of degree  $m = 3$  is given by  $N = 2$  paths

$$\omega_1(s) = -\frac{s}{\sqrt{t}}, \quad \omega_2(s) = \frac{s}{\sqrt{t}}$$

for fixed time horizon  $t$ . The corresponding weights are given by  $\lambda_1 = \lambda_2 = \frac{1}{2}$ .

Combining the stochastic Taylor expansion, the deterministic Taylor expansion for solutions of ODEs on  $H$  and a cubature formula on Wiener space yields a one-step scheme for weak approximation of the SDE (1) for bounded operators  $A$  in precisely the same way as in [18]. Indeed, we get

$$(11) \quad \sup_{x \in H} \left| E(f(X_t^x)) - \sum_{l=1}^N \lambda_l f(X_t^x(\omega_l)) \right| \leq C t^{\frac{m+1}{2}} \max_{\substack{(i_1, \dots, i_k) \in \mathcal{A} \\ m < \deg(i_1, \dots, i_k) \leq m+2}} \sup_{y \in H} |\beta_{i_1} \cdots \beta_{i_k} f(y)|,$$

for  $0 < t < 1$ ,  $f \in C^\infty(H)$ .

For the multi-step method, divide the interval  $[0, T]$  into  $p$  subintervals according to the partition  $0 = t_0 < t_1 < \dots < t_p = T$ . For a multi-index  $(l_1, \dots, l_p) \in \{1, \dots, N\}^p$  consider the path  $\omega_{l_1, \dots, l_p}$  defined by concatenating the paths  $\omega_{l_1}, \dots, \omega_{l_p}$ , i.e.  $\omega_{l_1, \dots, l_p}(t) = \omega_{l_1}(t)$  for  $t \in [0, t_1]$  and

$$\omega_{l_1, \dots, l_p}(t) = \omega_{l_1, \dots, l_p}(t_{r-1}) + \omega_{l_r}(t - t_{r-1})$$

for  $r$  such that  $t \in [t_{r-1}, t_r]$ , where  $\omega_{l_r}$  is scaled to be a cubature path on the interval  $[0, t_r - t_{r-1}]$ .

**Proposition 3.6.** Fix  $T > 0$ ,  $m \in \mathbb{N}$ , a cubature formula of degree  $m$  as in Definition 3.3 and a partition of  $[0, T]$  as above. For every  $f \in C^\infty(H)$  with

$$\sup_{0 \leq t \leq T} \sup_{y \in H} |\beta_{i_1} \cdots \beta_{i_k}(P_t f)(y)| < \infty$$

for all  $(i_1, \dots, i_k) \in \mathcal{A}$  with  $m < \deg(i_1, \dots, i_k) \leq m+2$ ,  $k \in \mathbb{N}$ , there is a constant  $D$  independent of the partition such that

$$\sup_{x \in H} \left| E(f(X_T^x)) - \sum_{(l_1, \dots, l_p) \in \{1, \dots, N\}^p} \lambda_{l_1} \cdots \lambda_{l_p} f(X_T^x(\omega_{l_1, \dots, l_p})) \right| \leq DT \max_{r=1, \dots, p} (t_r - t_{r-1})^{(m-1)/2}.$$

Note that the assumption on  $f$  in Proposition 3.6 is always satisfied if  $f$  is bounded,  $C^\infty$ -bounded and has bounded support and if all vector fields  $\alpha, \beta_1, \dots, \beta_d$  have bounded support (compare to Remark 3.2). In this case  $P_t f$  has bounded support, too, and all derivatives are bounded on the whole of  $H$  (see the discussion after Theorem 4.4 in the next section).

**Remark 3.7.** Proposition 3.6 also yields *deterministic a priori* estimates for the weak rate of convergence, which hold true if we are able to evaluate the respective tree deterministically (see also Section 6). In principle there are methods to do so, see [21], however, they might be more cumbersome to implement than a Monte-Carlo evaluation of the tree.

**Remark 3.8.** There is one general case where Proposition 3.6 can be applied: namely for stochastic differential equations of type (1). If we replace the unbounded operator  $A$  with a bounded operator  $\tilde{A}$ , which is close to  $A$  for a large enough set of values  $x \in \mathcal{D}(A)$ , then we can apply the previous result for bounded linear operators. One candidate for this procedure is the Yosida approximation of  $A$ .

#### 4. THE CASE WHEN $A$ IS UNBOUNDED LINEAR

We shall also assume in this section that there are no jumps and refer to equation (1). We recall the analytic background: let us denote by  $\mathcal{D}(A^k)$  the Hilbert space given by  $\mathcal{D}(A^k)$  equipped with the graph norm

$$\|h\|_{\mathcal{D}(A^k)}^2 = \sum_{i=0}^k \|A^i h\|^2,$$

$h \in \mathcal{D}(A^k)$  and  $k \geq 1$ . Furthermore, we introduce the space

$$\mathcal{D}(A^\infty) = \bigcap_{k=0}^{\infty} \mathcal{D}(A^k).$$

$\mathcal{D}(A^\infty)$  is topologized as projective limit of the Hilbert spaces  $\mathcal{D}(A^k)$ ,  $k \geq 0$ , i.e. the topology on  $\mathcal{D}(A^\infty)$  is the initial topology of the maps  $\mathcal{D}(A^\infty) \rightarrow \mathcal{D}(A^k)$ ,  $k \in \mathbb{N}$ . Note that  $\mathcal{D}(A^\infty)$  is no longer a Hilbert space, but only a Fréchet space, i.e. a locally convex vector space which is completely metrizable by a translation invariant metric. We assume our main Assumption 2.1, i.e. the vector fields restricted to the Sobolev spaces are  $C^\infty$ -bounded.

The following proposition collects a few easy, but interesting corollaries from the existence and uniqueness theorem for equation (1) applied to the situation specified in Assumption 2.1.

**Proposition 4.1.** Fix  $k \in \mathbb{N}$ . For any  $x \in \mathcal{D}(A^k)$  there is a unique continuous mild solution of (8) interpreted as an SDE in the Hilbert space  $\mathcal{D}(A^k)$ . If  $x \in \mathcal{D}(A^{k+1}) \subset \mathcal{D}(A^k)$ , then the mild solution in  $\mathcal{D}(A^k)$  coincides with the mild solution



in  $\mathcal{D}(A^{k+1})$ . Consequently, for  $x \in \mathcal{D}(A^{k+1})$ , the mild solution of (1) in  $\mathcal{D}(A^k)$  is a strong solution and, in particular, a semi-martingale.

If we start in  $x \in \mathcal{D}(A^\infty)$ , then we get a continuous process  $X_t^x \in \mathcal{D}(A^\infty)$ , such that  $X_t^x$  is the (mild and strong) solution of (8) in any Hilbert space  $\mathcal{D}(A^k)$ ,  $k \in \mathbb{N}$ . Furthermore, Proposition 4.1 allows us to avoid any problems due to the topological structure of  $\mathcal{D}(A^\infty)$  by reverting to an appropriate Hilbert space  $\mathcal{D}(A^k)$  and interpreting the results again in  $\mathcal{D}(A^\infty)$ . The meaning of  $\mathcal{D}(A^\infty)$  is that it is the largest subspace of the Hilbert space  $H$ , where we can innocently do the necessary analysis on differential operators. Notice that there are subtle phenomena of explosion, which can occur in this setting: for instance the law of a strong solution process  $X$  solving equation (1) might be bounded in  $H$  but unbounded in  $\mathcal{D}(A)$ , where it is a mild solution. Due to such phenomena, Example 4.8 and Remark 4.9 after Theorem 4.4 are in fact quite subtle.

As in Section 3, we introduce the vector field  $\beta_0$  defined by

$$(12) \quad \beta_0(x) = Ax + \alpha(x) - \frac{1}{2} \sum_{i=1}^d D\beta_i(x) \cdot \beta_i(x).$$

$\beta_0$  is defined for  $x \in \mathcal{D}(A)$ . As a vector field taking values in  $\mathcal{D}(A^k)$ , it is only well-defined on  $\mathcal{D}(A^{k+1})$ . Consequently, for  $x \in \mathcal{D}(A^{k+1})$ , we may reformulate the SDE (1) – understood as equation in  $\mathcal{D}(A^k)$  – in Stratonovich form (9).

Now we formulate the stochastic Taylor expansion in some  $\mathcal{D}(A^{r(m)})$  with a degree of regularity  $r(m) \geq 0$  depending on  $m \geq 1$ . For the estimation of the error term, we will use the *extended support*  $\text{esupp}(X_t^x; \omega_1, \dots, \omega_N)$  defined by

$$(13) \quad \text{esupp}(X_t^x; \omega_1, \dots, \omega_N) = \text{supp}(X_t^x) \cup \{X_t^x(\omega_1), \dots, X_t^x(\omega_N)\},$$

where  $t > 0$ ,  $x \in H$ , and  $\omega_1, \dots, \omega_N$  are paths of bounded variation. Here  $\text{supp}(X_t^x)$  means the support of the law of  $X_t^x$  in  $\mathcal{D}(A^{r(m)})$ . Despite Assumption 3.4, let us, for one moment, enter the dependence of the cubature formula on the interval  $[0, t]$  explicitly into the notation, in the sense that  $\omega_1^{(t)}, \dots, \omega_N^{(t)}$  are the paths of bounded variation scaled in such a way that they, together with the weights, form a cubature formula on  $[0, t]$ . Then we denote

$$\mathfrak{S}_T(x) = \bigcup_{0 \leq s \leq t \leq T} \text{esupp}(X_s^x; \omega_1^{(t)}, \dots, \omega_r^{(t)}).$$

**Remark 4.2.** If a general support theorem holds in infinite dimensions, we can replace the extended support by the ordinary support of  $X_t^x$ , since the solution of the corresponding ODE driven by paths of bounded variation lie in the support of the solution of the SDE according to the support theorem. However, up to our knowledge, no general support theorem has been established for our setting so far.

**Theorem 4.3.** *Let  $m \geq 1$  be fixed, then there is  $r(m) \geq 0$  such that for any  $f \in C^\infty(H; \mathbb{R})$ ,  $x \in \mathcal{D}(A^{r(m)})$ , and  $0 < t < 1$  we have*

$$f(X_t^x) = \sum_{\substack{k \leq m, (i_1, \dots, i_k) \in \mathcal{A} \\ \deg(i_1, \dots, i_k) \leq m}} (\beta_{i_1} \cdots \beta_{i_k} f)(x) B_t^{(i_1, \dots, i_k)} + R_m(t, f, x), \quad x \in \mathcal{D}(A^{r(m)}),$$

with

$$\begin{aligned} \sqrt{E(R_m(t, f, x)^2)} &\leq Ct^{\frac{m+1}{2}} \max_{m < \deg(i_1, \dots, i_k) \leq m+2} \sup_{0 \leq s \leq t} |E(\beta_{i_1} \cdots \beta_{i_k} f(X_s^x))| \\ &\leq Ct^{\frac{m+1}{2}} \max_{m < \deg(i_1, \dots, i_k) \leq m+2} \sup_{y \in \mathfrak{S}_t(x)} |\beta_{i_1} \cdots \beta_{i_k} f(y)|. \end{aligned}$$

We can choose  $r(m) = \lfloor \frac{m}{2} \rfloor + 1$ , where  $\lfloor \frac{m}{2} \rfloor$  is the largest integer smaller than  $\frac{m}{2}$ .

*Proof.* The proof is the same as in the finite dimensional situation, but one has to switch between different spaces on the way.

Fix  $m$  and  $f$  as above and  $x \in \mathcal{D}(A^{\lfloor \frac{m}{2} \rfloor + 1})$ . We interpret the equation in  $\mathcal{D}(A^{\lfloor \frac{m}{2} \rfloor + 1})$ . By the above remarks, we can express the SDE in its Stratonovich form (9). By Itô's formula,

$$(14) \quad f(X_t^x) = f(x) + \int_0^t (\beta_0 f)(X_s^x) ds + \sum_{i=1}^d \int_0^t (\beta_i f)(X_s^x) \circ dB_s^i.$$

The idea is to express  $(\beta_i f)(X_s^x)$  again by Itô's formula and insert it in equation (14). This is completely unproblematic for  $i \in \{1, \dots, d\}$ . For  $i = 0$ , recall that

$$(\beta_0 f)(x) = Df(x) \cdot Ax + Df(x) \cdot \left( \alpha(x) - \frac{1}{2} \sum_{i=1}^d D\beta_i(x) \cdot \beta_i(x) \right).$$

By re-expressing (14) in Itô formulation, applying Itô's formula, and re-expressing it back to Stratonovich formulation, we see that

$$(\beta_0 f)(X_s^x) = (\beta_0 f)(x) + \int_0^s (\beta_0^2 f)(X_u^x) du + \sum_{i=1}^d \int_0^s (\beta_i \beta_0 f)(X_u^x) \circ dB_u^i,$$

where

$$(15) \quad (\beta_0^2 f)(x) = D^2 f(x)(Ax, Ax) + Df(x) \cdot (A^2 x + A\alpha(x) + \dots) + \dots,$$

provided that all the new vector-fields are well-defined and the processes  $(\beta_i \beta_0 f)(X_u^x)$  are still semi-martingales. Both conditions are satisfied if  $x \in \mathcal{D}(A^2)$  – notice that the maps  $\mathcal{D}(A^{k+1}) \rightarrow \mathcal{D}(A^k)$ ,  $x \mapsto Ax$  are  $C^\infty$ ,  $k \in \mathbb{N}$ . By induction, we finally get

$$f(X_t^x) = \sum_{\substack{(i_1, \dots, i_k) \in \mathcal{A} \\ \deg(i_1, \dots, i_k) \leq m}} (\beta_{i_1} \cdots \beta_{i_k} f)(x) B_t^{(i_1, \dots, i_k)} + R_m(t, f, x)$$

with

$$\begin{aligned} & R_m(t, x, f) \\ &= \sum_{\substack{(i_1, \dots, i_k) \in \mathcal{A}, i_0 \in \{0, \dots, d\} \\ \deg(i_1, \dots, i_k) \leq m < \deg(i_0, i_1, \dots, i_k)}} \int_{0 \leq t_0 \leq \dots \leq t_k \leq t} (\beta_{i_0} \cdots \beta_{i_k} f)(X_{t_0}^x) \circ dB_{t_0}^{i_0} \cdots \circ dB_{t_k}^{i_k}. \end{aligned}$$

Note that  $R_m$  is well-defined for  $x \in \mathcal{D}(A^{\lfloor \frac{m}{2} \rfloor + 1})$  because integration of non-semi-martingales with respect to  $dt$  is possible, which corresponds to the index  $i_0 = 0$ .

As in the finite dimensional case, we re-express  $R_m$  in terms of Itô integrals and use the (one-dimensional) Itô isometry several times, until we arrive at the desired estimate.  $\square$

We recall the notation  $P_t f(x) = E(f(X_t^x))$  for bounded measurable functions  $f : H \rightarrow \mathbb{R}$ . Analogously to Proposition 3.6, we immediately get the following theorem.

**Theorem 4.4.** *Fix  $T > 0$ ,  $m \geq 1$ ,  $r(m)$ ,  $x \in \mathcal{D}(A^{r(m)})$  as in Theorem 4.3, a cubature formula on Wiener space of degree  $m$  as in Definition 3.3 and a partition  $0 = t_0 < t_1 < \dots < t_p = T$ . Under Assumption 2.1, for any  $f \in C^\infty(H; \mathbb{R})$  with*

$$\sup_{0 \leq t \leq T} \sup_{y \in \mathfrak{S}_T(x)} |\beta_{i_1} \cdots \beta_{i_k} P_t f(y)| < \infty$$

for all  $(i_1, \dots, i_k) \in \mathcal{A}$  with  $m < \deg(i_1, \dots, i_k) \leq m+2$ ,  $k \in \mathbb{N}$ , there is a constant  $D$  independent of the partition such that

$$\left| E(f(X_T^x)) - \sum_{(l_1, \dots, l_p) \in \{1, \dots, N\}^p} \lambda_{l_1} \cdots \lambda_{l_p} f(X_T^x(\omega_{l_1, \dots, l_p})) \right| \leq DT \max_{r=1, \dots, p} (t_r - t_{r-1})^{(m-1)/2},$$

where  $X_T^x(\omega)$  is, again, understood as the mild solution to an ODE in  $\mathcal{D}(A^{r(m)})$  for any path  $\omega$  of bounded variation.

*Proof.* The proof follows Kusuoka [16], [17], see also [14]. For  $f : H \rightarrow \mathbb{R}$  and  $x \in H$  let

$$Q_{(t)}f(x) = \sum_{l=1}^N \lambda_l f(X_t^x(\omega_l)),$$

where  $\omega_1, \dots, \omega_l$  are scaled to form a cubature formula on  $[0, t]$ . Denote  $\Delta t_r = t_r - t_{r-1}$ ,  $r = 1, \dots, p$ , the increments of the time partition given in the statement of the theorem. By iterating the operators  $Q_{(\Delta t_r)}$  (and the semigroup property of ODEs), we immediately obtain

$$(16) \quad \sum_{(l_1, \dots, l_p) \in \{1, \dots, N\}^p} \lambda_{l_1} \cdots \lambda_{l_p} f(X_T^x(\omega_{l_1, \dots, l_p})) = Q_{(\Delta t_p)} \circ \cdots \circ Q_{(\Delta t_1)} f(x).$$

By ordinary Taylor expansion, keeping in mind the degree function  $\deg$ , we note that

$$Q_{(t)}f(x) = \sum_{l=1}^N \lambda_l \sum_{\substack{k \leq m, (i_1, \dots, i_k) \in \mathcal{A} \\ \deg(i_1, \dots, i_k) \leq m}} (\beta_{i_1} \cdots \beta_{i_k} f)(x) B_t^{(i_1, \dots, i_k)}(\omega_l) + \tilde{R}_m(t, x, f),$$

where  $x \in \mathcal{D}(A^{r(m)})$  and

$$\begin{aligned} \tilde{R}_m(t, x, f) &= \sum_{l=1}^N \lambda_l \sum_{\substack{(i_1, \dots, i_k) \in \mathcal{A}, i_0 \in \{0, \dots, d\} \\ \deg(i_1, \dots, i_k) \leq m < \deg(i_0, \dots, i_k)}} \\ &\quad \int_{0 \leq t_0 \leq \dots \leq t_k \leq t} (\beta_{i_0} \cdots \beta_{i_k} f)(X_{t_0}^x(\omega_l)) d\omega_l^{i_0}(t_0) \cdots d\omega_l^{i_k}(t_k). \end{aligned}$$

In the sequel  $C$  denotes a constant independent of the partition and  $x$ , which may change from line to line. We can estimate the approximation error by

$$\left| \tilde{R}_m(t, x, f) \right| \leq C \sup_{0 \leq s \leq t, l=1, \dots, N} \max_{m \leq \deg(i_0, \dots, i_k) \leq m+2} |\beta_{i_0} \cdots \beta_{i_k} f(X_s^x(\omega_l))| t^{\frac{m+1}{2}}.$$

Combining this result with Theorem 4.3, we may conclude that

$$(17) \quad |P_t f(x) - Q_{(t)} f(x)| \leq C \sup_{y \in \mathfrak{S}_t(x)} |\beta_{i_0} \cdots \beta_{i_k} f(y)| t^{\frac{m+1}{2}}.$$

By telescopic sums,

$$\begin{aligned} P_T f(x) - Q_{(\Delta t_p)} \circ \cdots \circ Q_{(\Delta t_1)} f(x) &= \\ &= \sum_{r=1}^p Q_{(\Delta t_p)} \circ \cdots \circ Q_{(\Delta t_{r+1})} (P_{t_r} f(x) - Q_{(\Delta t_r)} P_{t_{r-1}} f(x)). \end{aligned}$$

For the estimation of the rear term

$$P_{t_r} f(x) - Q_{(\Delta t_r)} P_{t_{r-1}} f(x) = (P_{\Delta t_r} - Q_{(\Delta t_r)}) P_{t_{r-1}} f(x),$$

we may use (17) with  $f(x)$  being replaced by  $P_{t_{r-1}}f(x)$ , giving us

$$\begin{aligned}
|P_T f(x) - Q_{(\Delta t_p)} \circ \cdots \circ Q_{(\Delta t_1)} f(x)| &\leq \sum_{r=1}^p e^{CT} |P_{t_r} f(x) - Q_{(\Delta t_r)} P_{t_{r-1}} f(x)| \\
&\leq C \sum_{r=1}^p \sup_{\substack{y \in \mathfrak{S}_{\Delta t_r}(x) \\ m \leq \deg(i_0, \dots, i_k) \leq m+2}} |\beta_{i_0} \cdots \beta_{i_k} P_{t_{r-1}} f(y)| (\Delta t_r)^{\frac{m+1}{2}} \\
&\leq C \sup_{\substack{y \in \mathfrak{S}_T(x), 0 \leq t \leq T \\ m \leq \deg(i_0, \dots, i_k) \leq m+2}} |\beta_{i_0} \cdots \beta_{i_k} P_t f(y)| \sum_{r=1}^p (\Delta t_r)^{\frac{m+1}{2}},
\end{aligned}$$

from which we may easily conclude the theorem.  $\square$

**Remark 4.5.** Gyöngy and Shmatkov [11] show a strong Wong-Zakai-type approximation result, where they also need to impose smoothness assumptions on the initial value  $x$ . Otherwise, the assumptions in [11] are different from ours. They allow linear, densely defined vector fields and general adapted coefficients, on the other hand the generator  $A$  needs to be elliptic.

**Remark 4.6.** Under the previous assumptions we can also prove a Donsker-type result on the weak convergence of the “cubature tree” to the diffusion. This result will be presented elsewhere.

**Remark 4.7.** If  $f$  is smooth then we can show by (first and higher) variation processes, as introduced for instance in [8], that  $x \mapsto P_t f(x)$  is smooth on  $\mathcal{D}(A^k)$ .

Fix  $k \geq 0$ . Let  $J_{0 \rightarrow t}(x) \cdot h$  denote the first variation process of  $X_t^x$  in direction  $h \in \mathcal{D}(A^k)$ , i.e.

$$J_{0 \rightarrow t}(x) \cdot h = \left. \frac{\partial}{\partial \epsilon} \right|_{\epsilon=0} X_t^{x+\epsilon h} \in \mathcal{D}(A^k).$$

$J_{0 \rightarrow t}(x) \cdot h$  is the mild solution to an SDE of the type (8). Consequently, it is bounded in  $L^2(\Omega, \mathcal{F}, P; \mathcal{D}(A^k))$  and we may conclude that

$$\left. \frac{\partial}{\partial \epsilon} \right|_{\epsilon=0} P_t f(x + \epsilon h) = E \left( Df(X_t^x) \cdot J_{0 \rightarrow t}(x) \cdot h \right)$$

exists and is bounded by boundedness of  $Df$  and integrability of the first variation. Similarly, we get existence and continuity of higher order derivatives on  $\mathcal{D}(A^k)$ .

**Example 4.8.** We shall provide examples, where the assumptions of Theorem 4.4 are satisfied, i.e. where we obtain high-order convergence of the respective cubature methods. The conditions seem at first sight restrictive (see the following Remark 4.9 for a concrete example under Assumptions 2.1), however, the conditions are parallel to those obtained in [17] and [18], where the functions and vector fields have to be bounded and  $C^\infty$ -bounded. Here we have an additional complication of one certainly unbounded, but not even continuous drift vector field, which leads to the following set of assumptions.

The vector fields  $\alpha, \beta_1, \dots, \beta_d$  have the following property (compare also to tame maps in [12]): there exist smooth maps

$$\tilde{\alpha}, \tilde{\beta}_1, \dots, \tilde{\beta}_d : \mathcal{D}(A^{r(m)}) \rightarrow \mathcal{D}(A^{2r(m)})$$

such that for  $k \geq 0$  the restrictions to the respective subspaces  $\mathcal{D}(A^{k+r(m)}) \subset \mathcal{D}(A^{r(m)})$  take values in  $\mathcal{D}(A^{k+2r(m)})$ , i. e.

$$\tilde{\alpha}, \tilde{\beta}_1, \dots, \tilde{\beta}_d : \mathcal{D}(A^{k+r(m)}) \rightarrow \mathcal{D}(A^{k+2r(m)}),$$

and such that

$$(18) \quad \alpha = \tilde{\alpha} \circ (R(\lambda, A)^{r(m)}), \quad \beta_i = \tilde{\beta}_i \circ (R(\lambda, A)^{r(m)}),$$

for  $i = 1, \dots, d$ . Here  $R(\lambda, A)$  denotes the resolvent map for  $\lambda \in \rho(A)$ . We assume that  $\tilde{\alpha}, \tilde{\beta}_1, \dots, \tilde{\beta}_d$  have bounded support on  $\mathcal{D}(A^{r(m)})$ . The function  $f$  is of the same type  $f = g \circ (R(\lambda, A)^{r(m)})$  for a bounded,  $C^\infty$ -bounded function  $g : H \rightarrow \mathbb{R}$ .

Under these assumptions we can readily check that the law of the mild solution  $X_t^x$  starting at the initial value  $x \in \mathcal{D}(A^{r(m)})$  has bounded support in  $H$ : outside a ball of radius  $R > 0$  in  $H$  the solution process becomes deterministic,  $X_t^x = S_t x$  on some interval, hence by the uniform boundedness theorem there is a large  $R'$  such that the image of the ball with radius  $R > 0$  under the maps  $S_t$  lies in a ball with radius  $R' > 0$  on  $[0, T]$ .

For smooth functions  $f$  of the stated type we then have

$$\sup_{0 \leq t \leq T} \sup_{y \in H, \|y\| \leq R'} |\beta_{i_0} \cdots \beta_{i_k} P_t f(y)| < \infty.$$

Since we only take the supremum over bounded sets, namely the extended supports of  $X_t^x$ , this implies the assumption of the Theorem 4.4.

**Remark 4.9.** The previous assumptions (18) on the vector fields are not too restrictive since we can always obtain them by a linear isomorphism and (smoothly) cutting off outside a ball in  $\mathcal{D}(A^{r(m)})$ . Both operations are numerically innocent. Under Assumption 2.1 we can apply the following isomorphism to the solution of our SDE (1):

$$R(\lambda, A)^{-r(m)} : \mathcal{D}(A^{r(m)}) \rightarrow H.$$

This isomorphism transforms the solution  $X_t^x$  on  $\mathcal{D}(A^{r(m)})$  to an  $H$ -valued process

$$Y_t^y = R(\lambda, A)^{-r(m)} X_t^{R(\lambda, A)^{r(m)} y}$$

with  $x = R(\lambda, A)^{r(m)} y$ . The transformed process  $Y_t^y$  satisfies an SDE, where the transformed vector fields (if well defined) factor over  $R(\lambda, A)^{r(m)}$  such as previously assumed in the assumption (18), namely

$$(19) \quad dY_t^y = (AY_t^y + ((R(\lambda, A)^{-r(m)} \circ \alpha \circ R(\lambda, A)^{r(m)})(Y_t^y))dt + \\ + \sum_{i=1}^d (R(\lambda, A)^{-r(m)} \circ \beta_i \circ R(\lambda, A)^{r(m)})(Y_t^y)dB_t^i.$$

The assumptions (18) mean that we must (smoothly) cut off the vector fields  $\alpha, \beta_1, \dots, \beta_d$  outside sets of large norm  $\|\cdot\|_{\mathcal{D}(A^{r(m)})}$ , which is an event – under Assumption 2.1 – of small probability (recall that the vector fields are Lipschitz on  $\mathcal{D}(A^{r(m)})$  and therefore second moments with respect to the norm  $\|\cdot\|_{\mathcal{D}(A^{r(m)})}$  exist). Notice that the cut-off vector fields do not have an extension to  $H$  since continuous functions with bounded support on  $\mathcal{D}(A^{r(m)})$  do generically not have a continuous extension on  $H$ . For  $(Y_t^y)_{t \geq 0}$  we can take initial values  $y \in H$ , however, those initial values correspond to quite regular initial values  $x = R(\lambda, A)^{r(m)} y \in \mathcal{D}(A^{r(m)})$  for the original process  $(X_t^x)_{t \geq 0}$ .

From the point of view of the process  $Y$  we have hence proved that

$$f(Y_T^y) = g \circ (R(\lambda, A)^{r(m)}) \circ (R(\lambda, A)^{-r(m)})(X_T^x) = g(X_T^x)$$

is weakly approximated by evaluating  $f$  on the cubature tree for  $Y$ . This is equivalent to evaluating  $g$  on the cubature tree for  $X$  in order to approximate the expected value  $E(g(X_T^x))$ .

## 5. THE CUBATURE METHOD IN THE PRESENCE OF JUMPS

The extension of cubature formulas to jump diffusions seems to be new even in the finite dimensional case. We shall heavily use the fact that only finitely many jumps occur in compact time intervals almost surely.

We shall first prove an asymptotic result on jump-diffusions.

**Proposition 5.1.** *Consider equation (2). Let  $f : H \rightarrow \mathbb{R}$  be a bounded measurable function, then we obtain*

$$(20) \quad E(f(X_t^x)) = \sum_{n_1, \dots, n_e \geq 0} \frac{\mu_1^{n_1} \dots \mu_e^{n_e}}{n_1! \dots n_e!} e^{-t\mu_1 n_1 - \dots - t\mu_e n_e} t^{n_1 + \dots + n_e} \times \\ \times E(f(X_t^x) \mid N_t^j = n_j \text{ for } j = 1, \dots, e)$$

for  $t \geq 0$ .

*Proof.* We condition on the jump times and read off the results by inserting the probabilities for a Poisson process with intensity  $\mu_j$  to reach level  $n_j$  at time  $t$ .  $\square$

This result gives us the time-asymptotics with respect to the jump-structure. It can now be combined with the original cubature result for the diffusion between the jumps, in order to obtain a result for jump-diffusions. We denote by  $\tau_n^j$  the jump-time of the Poisson process  $N^j$  for the  $n$ -th jump. We know that for each Poisson process the vector  $(\tau_1^j, \dots, \tau_{n_j}^j - \tau_{n_j-1}^j, t - \tau_{n_j}^j)$  is uniformly distributed if conditioned on the event that  $N_t^j = n_j \geq 1$ . The uniform distribution is on the  $n_j$ -simplex  $t\Delta^{n_j} \subset \mathbb{R}^{n_j+1}$ . This allows us to apply an original cubature formula between two jumps of order  $m - 2n_1 - \dots - 2n_e$ , since we gain, for each jump, one order of time-asymptotics from the jump structure.

Assume now that the jump distributions  $\nu_j$  are concentrated at one point  $z_j \neq 0$ , i.e.  $\Delta L_{\tau_k^j}^j = z_j$  for  $j = 1, \dots, e$  and  $k \geq 1$ . If we want to consider a general jump-structure this amounts to an additional integration with respect to the jump distribution  $\nu_j$ .

Now we define an short-time approximation for the conditional expectations

$$E(f(X_t^x) \mid N_t^j = n_j \text{ for } j = 1, \dots, e)$$

of order  $m - 2n_1 - \dots - 2n_e$  with  $n_1 + \dots + n_e \leq \frac{m+1}{2}$ . Expressed in words, we are going to do the following: starting from the initial value  $x \in \mathcal{D}(A^{r(m)})$  we solve the stochastic differential equation (2) along the cubature paths  $\omega_l$  with probability  $\lambda_l > 0$ ,  $l = 1, \dots, N$  until the first jump appears. We collect the end-points of the trajectories, add the jump size at these points and start a new cubature method from the resulting points on. Notice that we can take a cubature method of considerably lower degree since every jump increases the local order of time-asymptotics by 1. The jump times are chosen independent and uniformly distributed on simplices of certain dimension  $n_j$  such that  $n_1 + \dots + n_e = n \leq \frac{m+1}{2}$ . We denote the cubature trajectory between jump  $\tau_{q-1}^j$  and  $\tau_q^j$  for  $1 \leq q \leq n_j$  with  $\omega_{l,j,q}$ . If  $n_j = 0$  no trajectories are associated. Hence we obtain the following theorem:

**Theorem 5.2.** *Fix  $m \geq 1$ . Consider the stochastic differential equation (2) under the condition  $N_t^j = n_j$  for  $j = 1, \dots, m$  with  $n_1 + \dots + n_e = n$  along concatenated trajectories of type  $\omega_{l,j,q}$ . Choose a cubature method of degree*

$$m' = m - 2n \geq 1.$$

*Concatenation is only performed with increasing  $q$ -index and a typical concatenated trajectory is denoted by  $\omega_{l_1, \dots, l_n}$ . Here we have in mind that the intervals, where*

the chosen path is  $\omega_{l,j,q}$ , come from a jump of  $N^j$  and have length  $\tau_q^j - \tau_{q-1}^j$ . Then there is  $r(m') \geq 0$  such that

$$(21) \quad |E(f(X_t^x) \mid N_t^j = n_j) - \sum_{l_1, \dots, l_n=1}^N \lambda_{l_1} \dots \lambda_{l_n} E(f(X_t^x(\omega_{l_1, \dots, l_n})) \mid N_t^j = n_j)| \\ \leq Ct^{\frac{m'+1}{2}} \max_{\substack{(i_1, \dots, i_k) \in \mathcal{A} \\ \deg(i_1, \dots, i_k) \leq m+2}} \sup_{\substack{y \in \text{supp}(X_s^x), \\ 0 \leq s \leq t}} |\beta_{i_1} \dots \beta_{i_k} E(f(X_{\tau_q, t}^y(\omega_{l_{q+1}, \dots, l_n})) \mid N_t^j = n_j)|,$$

where  $X_t^x(\omega)$  means the solution of the stochastic differential equation (2) in Stratonovich form

$$(22) \quad dX_t^x(\omega) = \beta_0(X_{t-}^x(\omega))dt + \sum_{i=1}^d \beta_i(X_{t-}^x(\omega)) \circ dB_t^i + \sum_{j=1}^e \delta_j(X_{t-}^x) dL_t^j,$$

along the trajectory  $\omega$ .

*Proof.* By our main Assumption 2.1 we know that the linkage operators  $x \mapsto \delta^j(x)$  are  $C^\infty$ -bounded on each  $\mathcal{D}(A^k)$ , hence through concatenation the errors, which appear on each subinterval  $[t_{n-q}, t]$  are of the type  $y \mapsto E(X_{\tau_q, t}^y(\omega_{l_{q+1}q, \dots, l_n}))$  for some  $1 \leq q \leq n$ . Taking the supremum yields the result.  $\square$

Combining the previous result with Proposition 5.1 yields under certain conditions on the vector fields (see the discussion after Theorem 4.3 in the previous section) by the triangle inequality that there is a constant  $D > 0$  such that

$$(23) \quad |E(f(X_t^x)) - \sum_{2(n_1 + \dots + n_e) \leq m} \sum_{l_1, \dots, l_n=1}^N \frac{\mu_1^{n_1} \dots \mu_e^{n_e}}{n_1! \dots n_e!} e^{-t\mu_1 n_1 - \dots - t\mu_e n_e} \times \\ \times \lambda_{l_1} \dots \lambda_{l_n} E(f(X_t^x(\omega_{l_1, \dots, l_n})) \mid N_t^j = n_j)| \leq Dt^{\frac{m+1}{2}} < \infty,$$

By iteration of the previous result we obtain in precisely the same manner as in Section 4 a cubature method of order  $m$  by applying several cubature methods of order  $m' \leq m$  between the jumps.

**Remark 5.3.** The only random element in the expectation  $E(f(X_t^x(\omega_{l_1, \dots, l_n})) \mid N_t^j = n_j)$  is given by the jump times  $\tau_k^j$ , which vary on certain simplices. For the implementation one has to simulate the uniform distributions on the simplices  $t\Delta^k$ . Since the integrals on the simplices  $t\Delta^k$  only have continuous integrands, we cannot hope for other methods than Monte-Carlo. The evaluation by a Monte-Carlo-algorithm can also be seen as a random choice of the concatenation grid for the constructed  $\omega_{l_1, \dots, l_n}$ . Another view could be to see a deterministic grid for the diffusion which is saturated by points where jumps occur. Implementation will be done elsewhere.

## 6. NUMERICAL EXAMPLES

Due to the previous results a numerical scheme for equation (1) can be set up by the following steps. Notice that we have a weak order of approximation  $\frac{m-1}{2}$  only under the assumptions of the previous section. In order to obtain those assumptions one has to modify a general equation of type (1) by smoothing procedures. These modifications can be done in a controlled way, more precisely, for each modification we have a rate of convergence to the un-modified object.

- Approximate the vector fields  $\alpha, \beta_1, \dots, \beta_d$  by vector fields satisfying the Assumptions 2.1. If the original vector fields are globally Lipschitz one can do this approximation with a rate of convergence for the  $L^2$ -distance of the original solution process and its approximation.

- Choose a degree of accuracy  $m \geq 2$ , which determines the weak order of convergence  $\frac{m-1}{2}$  in the sequel. Associated to  $m$  the number  $r(m)$  can be identified, which tells us about the degree of regularity of the initial value  $x \in H$ , which one needs for the assertions of Theorem 4.4.
- Identify due to the previous specifications a radius  $R > 0$  such that the  $\mathcal{D}(A^{r(m)})$ -norm of  $X_T^x$  is rarely beyond  $R$ . Cut-off the vector fields smoothly on  $\mathcal{D}(A^{r(m)})$  and verify assumptions (18) – maybe after smoothing – for the transformed process  $Y$  such as exercised in Remark 4.9.
- The resulting tree of trajectories yields a finite number of non-autonomous PDEs (5), which have to be evaluated. In the implementation one calculates with the smoothened vector fields satisfying the Assumptions 2.1, but not with the cut-off vector fields, since a large  $\mathcal{D}(A^{r(m)})$ -norm value  $R$  is reached with small probability due to its very choice.

We test the cubature method for two concrete examples: one toy example, where explicit solutions of the SPDE are readily available, and another, more interesting but still very easy example. Since cubature on Wiener space is a weak method, we calculate the expected value of a functional of the solution to the SPDE in both cases, i.e. the outputs of our computations are real numbers.

The results presented here are calculated in MATLAB using the built-in PDE-solver `pdepe` for solving the deterministic PDEs given by inserting the cubature paths into the SPDE under consideration. This PDE solver depends on a space grid given by the user as well as on a time grid, which is not very critical because it is adaptively refined by the program.

We do not use recombination techniques for cubature on Wiener space as in [21] and use the simplest possible cubature formula for  $d = 1$  Brownian motions:

$$\omega_1^{(T)}(t) = -\frac{t}{\sqrt{T}}, \quad \omega_2^{(T)} = \frac{t}{\sqrt{T}}, \quad t \in [0, T],$$

with weights  $\lambda_1 = \lambda_2 = \frac{1}{2}$  define a cubature formula of degree  $m = 3$  on  $[0, T]$ . Consequently, solving an SDE on a Hilbert space with cubature on Wiener space for the above cubature formula and  $p$  iterations means solving  $2^p$  PDEs. This starts to get restrictive even for a very simple problem for, say,  $p = 10$  – where one already has to solve more than one thousand PDEs. One possibility to overcome these tight limitations is to use “a Monte-Carlo simulation on the tree”. Recall that an  $p$ -step cubature method approximates

$$E(f(X_T^x)) \approx \sum_{(j_1, \dots, j_p) \in \{1, \dots, N\}^p} \lambda_{j_1} \cdots \lambda_{j_p} f(X_T^x(\omega_{j_1, \dots, j_p})).$$

Since  $\sum \lambda_{j_1} \cdots \lambda_{j_p} = 1$ , we can interpret the right hand side as the expectation of a random variable  $f(X_T^x(\omega))$  on the tree  $\{1, \dots, N\}^p$ . Therefore, we can approximate the right hand side by picking tuples  $(j_1, \dots, j_p) \in \{1, \dots, N\}^p$  at random – according to their probabilities  $\lambda_{j_1} \cdots \lambda_{j_p}$  – and calculating the average of the corresponding outcomes  $f(X_T^x(\omega_{j_1, \dots, j_p}))$ . Of course, by following this strategy we have to replace the deterministic error estimates by a stochastic one, which heavily depends on the standard deviation of  $f(X_T^x(\omega))$  understood as a random variable on the tree. Notice, however, that this is the usual situation for weak approximation methods for SDEs.

Consider the Ornstein-Uhlenbeck process  $X_t^x$  defined as solution to the equation

$$(24) \quad dX_t^x = \Delta X_t^x dt + \phi dB_t$$

on the Hilbert space  $H = L^2([0, 1])$ .  $\Delta$  denotes the Dirichlet Laplacian on  $]0, 1[$ , i.e.  $\Delta$  is a negative definite self-adjoint operator on  $H$  with  $\mathcal{D}(\Delta) = H_0^1([0, 1]) \cap H^2([0, 1])$  extending the classical Laplace operator defined on  $C_c^\infty([0, 1])$ . It is easy



to see that  $\Delta$  is dissipative and therefore, by the Lumer-Phillips theorem, it is the generator of a  $C_0$  contraction semigroup  $(S_t)_{t \geq 0}$  on  $H$ . The coefficient  $\phi \in H$  is some fixed vector.

In this case, the definition of a mild solution

$$(25) \quad X_t^x = S_t x + \int_0^t S_{t-s} \phi dB_s$$

already gives a representation of the solution provided that the heat-semigroup  $S_t$  applied to the starting vector  $x$  and to  $\phi$  is available. We choose  $x(u) = \sin(\pi u)$ ,  $u \in ]0, 1[$ , and may conclude that

$$S_t x = e^{-\pi^2 t} x$$

because  $x$  is an eigenvector of  $\Delta$  with eigenvalue  $-\pi^2$ . Consider the linear functional  $\Phi : H \rightarrow \mathbb{R}$  given by

$$(26) \quad \Phi(y) = \int_0^1 y(u) du, \quad y \in H.$$

We want to compute

$$\begin{aligned} E(\Phi(X_1^x)) &= E\left(\int_0^1 e^{-\pi^2} \sin(\pi u) du + \int_0^1 \int_0^1 S_{1-s} \phi(u) dB_s du\right) \\ &= \int_0^1 e^{-\pi^2} \sin(\pi u) du = 0.3293 \times 10^{-4}. \end{aligned}$$

$p$	Error
1	$-0.3601 \times 10^{-4}$
2	$-0.2192 \times 10^{-4}$
3	$-0.1226 \times 10^{-4}$
4	$-0.0652 \times 10^{-4}$
5	$-0.0334 \times 10^{-4}$
6	$-0.0172 \times 10^{-4}$
7	$-0.0084 \times 10^{-4}$
8	$-0.0031 \times 10^{-4}$
9	$-0.0002 \times 10^{-4}$
10	$-0.0013 \times 10^{-4}$

TABLE 1. Error for the cubature method in the OU-case (absolute error)

In Table 1, the error, i.e. the output of the method minus the true value given above, is presented.  $p$  is the number of cubature steps, i.e. the number of iterations of the one-step cubature method. The discretization in space, i.e. of  $]0, 1[$ , used by the PDE solver contains 50 uniform points, the discretization of the time-interval – additional to the one induced by the cubature method – contains 500 points. The stochastic perturbation factor  $\phi$  is chosen to be  $\phi(u) = \sin(\pi u)$ , i.e.  $\phi \in \mathcal{D}(\Delta^\infty)$  even. We see a very fast decrease of the error in this simple situation. On the other hand, the variance of the random variable on the tree considered before is clearly too high for the Monte-Carlo simulation on the tree to work. Indeed,  $\Phi(X_1^x)$  has true standard deviation of

$$(27) \quad \text{sd}(\Phi(X_1^x)) = \sqrt{\frac{2}{\pi^4} (1 - e^{-2\pi^2})} = 0.1433.$$

Assuming that the central limit theorem applies, confidence intervals around the solution given by a Monte-Carlo method are proportional to the standard deviation

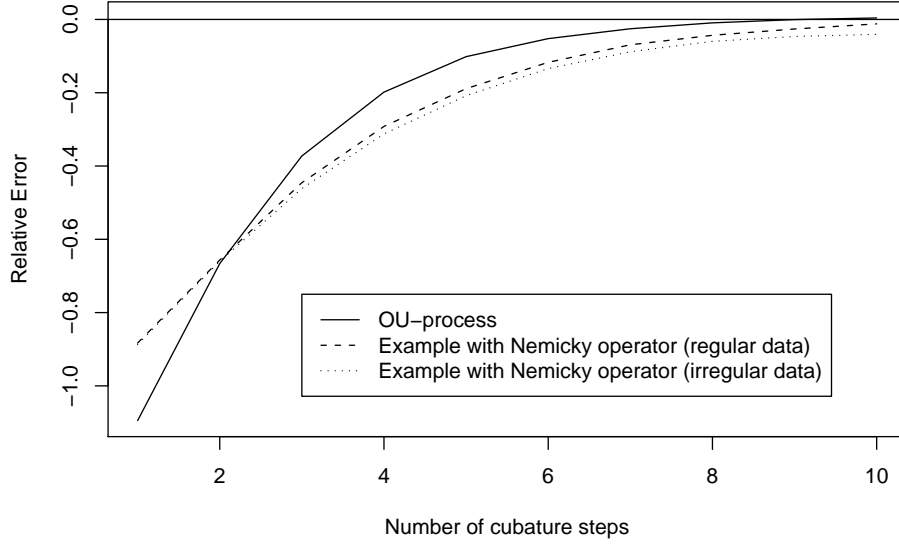


FIGURE 1. Relative errors for the OU-process (equation (24)) and the process with nonlinear volatility (equation (28)). In the latter case, both initial values  $x(u) = \sin(\pi u)$  – referred to as “regular case” – and  $x$  given by (31) – referred to as “irregular case” – are used.

divided by the square root of the number of trajectories. Consequently, we would roughly need to calculate  $10^{12}$  paths on the tree in order to achieve a similar level of exactness as in Table 1! Indeed, note that the standard deviation of the solution is of order  $10^{-1}$ , while the error in the last row of Table 1 is of order  $10^{-7}$ . The equation

$$\frac{10^{-1}}{\sqrt{m}} \approx 10^{-7}$$

then gives  $m \approx 10^{12}$ . Note that this heuristics is also confirmed by our experiments, where Monte-Carlo simulation on the tree clearly fails. The data are also shown in Figure 1.

**Remark 6.1.** The failure of Monte-Carlo simulation on the tree also applies to any other (naive) Monte-Carlo approach to problem (24), including the usual finite element or finite difference approaches.

As a more realistic example we consider the heat equation with a stochastic perturbation involving a Nemicky operator. More precisely, consider

$$(28) \quad dX_t^x = \Delta X_t^x dt + (\sin \circ X_t^x) dB_t,$$

with  $x(u) = \sin(\pi u)$ . Even though we do not know the law of the solution  $X_1^x$  of (28), we are still able to calculate  $E(\Phi(X_1^x))$  explicitly because  $\Phi$  is a linear functional. Indeed,  $X_1^x$  is given by

$$(29) \quad X_1^x = S_1 x + \int_0^1 S_{1-s} (\sin \circ X_s^x) dB_s$$

and, consequently,

$$\Phi(X_1^x) = \Phi(S_1 x) + \int_0^1 \Phi(S_{1-s} \sin \circ X_s^x) dB_s.$$

The expectation of the (one-dimensional) Itô-integral is 0 and we get the same result as before, i.e.

$$E(\Phi(X_1^x)) = \Phi(S_1 x) = 0.3293 \times 10^{-4}$$

for  $x(u) = \sin(\pi u)$ . Nevertheless, we believe that this example is already quite difficult, especially since the cubature method actually has to work with the Stratonovich formulation

$$(30) \quad dX_t^x(\omega) = \left( \Delta X_t^x(\omega) - \frac{1}{2} (\cos \circ X_t^x(\omega)) (\sin \circ X_t^x(\omega)) \right) dt + (\sin \circ X_t^x(\omega)) d\omega(t).$$

In particular, the equation (in Stratonovich form) has a non-linear drift and a non-linear volatility.

Note that we expect the standard deviation of the solution of the above equation to be smaller than before, because  $(\sin \circ X_t^x)^2$  decreases as  $X_t^x$  decreases in  $t$ . The

$l$	Error
1	$-0.2907 \times 10^{-4}$
2	$-0.2163 \times 10^{-4}$
3	$-0.1467 \times 10^{-4}$
4	$-0.0961 \times 10^{-4}$
5	$-0.0622 \times 10^{-4}$
6	$-0.0385 \times 10^{-4}$
7	$-0.0228 \times 10^{-4}$
8	$-0.0142 \times 10^{-4}$
9	$-0.0086 \times 10^{-4}$
10	$-0.0040 \times 10^{-4}$

TABLE 2. Results of the cubature method for (28) (absolute error)

space discretization used by the PDE-solver has size 50, which already seems to be sufficient, because using a finer discretization (100 grid points) does not change the results significantly. Table 3 shows the results using Monte-Carlo simulation

$l$	$m$	Error	Stat. Error
5	32	$0.0567 \times 10^{-4}$	$0.1498 \times 10^{-4}$
10	1000	$-0.0325 \times 10^{-4}$	$0.0179 \times 10^{-4}$
15	1500	$-0.0184 \times 10^{-4}$	$0.0172 \times 10^{-4}$
20	2000	$0.0128 \times 10^{-4}$	$0.0170 \times 10^{-4}$
25	2500	$0.0179 \times 10^{-4}$	$0.0145 \times 10^{-4}$
30	3000	$0.0596 \times 10^{-4}$	$0.0167 \times 10^{-4}$

TABLE 3. Results of the cubature method with Monte-Carlo simulation on the tree for (28) (absolute error)

on the tree.  $m$  denotes the number of trajectories followed, while the “Statistical Error” in the table is an indicator for the error of the Monte-Carlo simulation. More precisely, the values in the last column are the empirical standard deviations of the result divided by the square root of the number of trajectories. Comparable to the Ornstein-Uhlenbeck process, the convergence of the pure cubature method is very fast, see Table 2. The (empirical) variance is, however, quite large such that

the Monte-Carlo aided method does not work at all. Note that the statistical error in Table 3 is of the order of the total computational error, which can be almost completely attributed to the Monte Carlo simulation.

To test the method further we also try more irregular data. Let

$$(31) \quad x(u) = \frac{1}{2} \sqrt{\frac{1 - 2 \left| u - \frac{1}{2} \right|}{\sqrt{\left| u - \frac{1}{2} \right|}}}.$$

The exact value of the quantity of interest  $E(\Phi(X_1^x)) = \Phi(S_1 x)$  is calculated by solving the corresponding heat equation numerically. This gives the value  $E(\Phi(X_1^x)) = 0.3002 \times 10^{-4}$ . The initial vector  $x$  given in (31) is in  $L^2([0, 1])$  but its derivative is no longer square-integrable. Consequently,  $x \notin \mathcal{D}(A)$  and the theory does not provide an order of approximation. Nevertheless, probably due to the smoothing-properties of the Laplace operator numerical results show the same behavior as before, see Figure 1.

If we replace the heat equation (24) by an evolution equation of the form

$$(32) \quad dX_t^x = \frac{d}{du} X_t^x dt + \sin \circ X_t^x dB_t,$$

then we still see the same kind of behavior if we fix the space-discretization for the PDE-solver. This time, the PDEs require a much finer space resolution in order to give reliable numbers.

## REFERENCES

- [1] Fabrice Baudoin. *An introduction to the geometry of stochastic flows*. Imperial College Press, London, 2004.
- [2] Fabrice Baudoin and Josef Teichmann. Hypocoellipticity in infinite dimensions and an application in interest rate theory. *Ann. Appl. Probab.*, 15(3):1765–1777, 2005.
- [3] Tomas Björk, Anders Szepessy, Raul Tempone, and Georgios Zouraris. Monte-Carlo Euler approximation of HJM term structure financial models. Working paper.
- [4] Rama Cont and Peter Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [5] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.
- [6] Damir Filipovic and Stefan Tappe. Existence of Lévy term structure models. Forthcoming in *Finance and Stochastics*, 2006.
- [7] Damir Filipović and Josef Teichmann. Existence of invariant manifolds for stochastic equations in infinite dimension. *J. Funct. Anal.*, 197(2):398–432, 2003.
- [8] Barbara Forster, Eva Lütkebohmert, and Josef Teichmann. Absolutely continuous laws of jump-diffusions in finite and infinite dimensions with applications to mathematical finance. Submitted preprint available at <http://arxiv.org/abs/math/0509016>.
- [9] István Gyöngy. Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I. *Potential Anal.*, 9(1):1–25, 1998.
- [10] István Gyöngy. Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II. *Potential Anal.*, 11(1):1–37, 1999.
- [11] István Gyöngy and Anton Shmatkov. Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. *Appl. Math. Optim.*, 54(3):315–341, 2006.
- [12] Richard S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.
- [13] Erika Hausenblas. Approximation for semilinear stochastic evolution equations. *Potential Anal.*, 18(2):141–186, 2003.
- [14] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.
- [15] Andreas Kriegl and Peter W. Michor. *The convenient setting of global analysis*, volume 53 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

- [16] Shigeo Kusuoka. Approximation of expectation of diffusion process and mathematical finance. In *Taniguchi Conference on Mathematics Nara '98*, volume 31 of *Adv. Stud. Pure Math.*, pages 147–165. Math. Soc. Japan, Tokyo, 2001.
- [17] Shigeo Kusuoka. Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. In *Advances in mathematical economics. Vol. 6*, volume 6 of *Adv. Math. Econ.*, pages 69–83. Springer, Tokyo, 2004.
- [18] Terry Lyons and Nicolas Victoir. Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):169–198, 2004.
- [19] S. Peszat and J. Zabczyk. *Stochastic partial differential equations with Lévy noise*, volume 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007. An evolution equation approach.
- [20] Claudia Prévôt and Michael Röckner. *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.
- [21] Christian Schmeiser, Alexander Soreff, and Josef Teichmann. Recombination of cubature trees for the weak solution of sdes. Working paper.
- [22] Josef Teichmann. Calculating the Greeks by cubature formulae. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 462(2066):647–670, 2006.
- [23] Yubin Yan. Galerkin finite element methods for stochastic parabolic partial differential equations. *SIAM J. Numer. Anal.*, 43(4):1363–1384 (electronic), 2005.

VIENNA UNIVERSITY OF TECHNOLOGY, DEPARTMENT OF MATHEMATICAL METHODS IN ECONOMICS, RESEARCH GROUP E105 FINANCIAL AND ACTUARIAL MATHEMATICS, WIEDNER HAUPTSTRASSE 8-10, A-1040 WIEN, AUSTRIA

*E-mail address:* cbayer@fam.tuwien.ac.at, jteichma@fam.tuwien.ac.at